Hamilton College
Russell Marcus

Class #18: October 27
Mathematical Truth

I. Setting Up the Benacerraf Problem

Benacerraf's paper presents a problem which shaped nearly all discussions of the philosophy of mathematics in the second half of the twentieth century.
In order to explain the Benacerraf problem, I will first discuss two more-general philosophical topics.
In setting up the problem, Benacerraf presents two conditions.
The first condition involves semantics: mathematical semantics should be consistent with our broader views on semantics.
Semantics, on Benacerraf's view, is not limited to theories of meaning, but involves theories of truth and reference, which might be considered metaphysical rather than purely semantic.
Benacerraf's view of semantics, though, is (or at least was at the time) standard.
Semantic theories were/are seen as including specifications of truth conditions for the sentences of a language.
Benacerraf relies on Tarski's theory of truth, which remains the standard view of truth in philosophy.

Benacerraf's second condition is that mathematical epistemology should be consistent with our broader views about epistemology.
Our accounts of mathematical knowledge should be part of our wider account of knowledge.
Benacerraf relies on a causal theory of knowledge as his preferred epistemology.
As Field shows, the causal theory is not essential to the Benacerraf problem.
But, since Benacerraf frames the problem in those terms, it will be useful to discuss it.

I will briefly contextualize and explain the causal theory of knowledge and (less-briefly) Tarski's theory of truth before we get into more detail about the Benacerraf problem.

II. 2500 Years of Epistemology in Ten Minutes

The standard definition of knowledge is justified true belief (JTB).

X knows that p iff X believes that p and X is justified in believing that p and p is true

The JTB definition traces back to Plato's characterization of knowledge as true belief with a logos attached; see *Theaetetus* 200-201.
For the most part, the standard definition was taken as settled until 1962, when Gettier published his counter-examples.
Gettier presents two examples of justified true beliefs that we would not want to call knowledge.
In the first case, Smith believes that the man who will get a job has ten coins in his pocket because he believes (on good evidence) that Jones will get the job and he also believes that Jones has ten coins in his pocket (because he counted them).
Smith does not know how many coins he has in his own pocket.
It turns out, though, that Smith himself gets the job, and that Smith has ten coins in his pocket.
Smith's belief is true, and justified.
But it is not knowledge.

In the second case, Smith believes that either Jones owns a Ford or Brown is in Barcelona, since he believes that Jones owns a Ford and he knows the logical rule of addition.
He does not know where Brown is, but he has good evidence that Jones owns a Ford.
It turns out, though, that Jones is just leasing his Ford, and Brown is, in fact, in Barcelona.
Again, Smith has a justified true belief, but not knowledge.

Since the Gettier cases show that people can have a JTB without knowledge, the JTB definition is insufficient for knowledge.
The Gettier counter-examples led to an explosion of interest in epistemology, which had traditionally focused on other matters, like skepticism and foundationalism.
Suddenly, epistemologists did not know what they were chasing.
In the years since Gettier's paper, additional conditions were proposed to fix the JTB account.

One attempt to define knowledge, in the wake of Gettier's demolition of JTB, used a causal theory (CTK).
CTK can be understood as adding a fourth condition on JTB: the justification has to include appropriate causal connections between the knower and the proposition known.
In the case of the man who will get the job, Smith does not have an appropriate causal connection to the object of his knowledge, which in this case is the coins in the pocket of Smith himself, rather than those in Jones's pocket.
CTK gets the answer right: Smith does not know that the man who will get the job has ten coins in his pocket.

Benacerraf, writing after Gettier and the development of the causal theory, invokes CTK in 'Mathematical Truth'.
Unfortunately, by the mid-1970s, it became clear that CTK was itself flawed.
Part of the problem was the obscurity of the notion of causation on which it depended.
Aside from some purely philosophical worries about distinguishing between causes and mere antecedents, like those that Hume raised, some results from physics undermined the confidence that defenders of CTK placed in causation.
Physicists began considering the possibilities of tachyons, particles which travel faster than light.
Suddenly, backwards causation, in which an effect temporally precedes its cause, seemed possible.
Tachyons traveling forward in time with respect to one set of reference frames could be seen as traveling backwards in time from another set of reference frames.
If our choices of reference frames are arbitrary, as is consistent with the theory of relativity, then there is no physical motivation to insist that causes always precede effects.

A different objection came from Alvin Goldman, who had himself contributed to the development of CTK.
Goldman's counter-example to CTK uses the fake barn country example.
In fake barn country, there are thousands of barn facades.
From the road, the facades look like real barns, but they are not real.
There are also one or two real barns, among the thousands of fake ones.
If you are, unknowingly, driving through fake barn country, and happen to see one of the rare real barns, you might believe that you have seen a barn.
You would have a JTB that you have seen a barn.
In addition, you would be appropriately causally connected to a barn.

So you would fulfil the extra condition arising from CTK.
But, since you would have been in the same belief state had you seen one of the fake barns, we can not really say that you know that you have seen a barn.

Here is a link to a good account of the history that I have sketched:
http://plato.stanford.edu/entries/knowledge-analysis/#GET

III. Theories of Truth: Inflationary and Deflationary (Or, 2500 Years of Truth in Fifteen Minutes)

Benacerraf argues that just as we should have a consistent epistemology across science and mathematics, we should have a consistent semantic theory.
Benacerraf relies on Tarski's theory of truth in explaining the requirements of semantic theory.
Semantic theories can have three parts: a theory of truth, a theory of reference, and a theory of meaning.
A term, like 'cat', has some meaning, and refers to certain objects.
A sentence, like 'the cat is on the mat', has some meaning, and some truth conditions.
It may also have a reference.
We won't spend time on meaning theories.
We are mainly concerned here with the theories of truth and reference.
Of course, the meaning of a term and its reference may be closely connected.
Frege, in fact, defined the meaning of a term as that which determines its reference.
Reference is also closely connected to truth, since a sentence, like 'the cat is on the mat' seems to require, for its truth, that the term 'cat' refer to some specific cat, and the predicate 'is on the mat' refer to some sort of property or relation of being on the mat.

Aristotle's claim about truth remains central to our concept.

> To say of what is that it is not, or of what is not that it is, is false, while to say of what is that it is, and of what is not that it is not, is true (*Metaphysics* 1011b25).

Aristotle's theory of truth is standardly taken as a correspondence theory.
A correspondence theory says that truth is a correspondence between words and worlds.
A proposition is true if the world is the way that the proposition claims that it is.

One worry about correspondence theories is that we do not seem to have any extra-linguistic way to apprehend reality.
If I want to compare, say, an elephant to a picture of an elephant, or a picture of a sculpture of an elephant to a picture of an elephant, I can hold both of them in front of me, gazing from the one to the other.

If I want to compare my words to the world, I have to apprehend, on the one side, what the words mean, and on the other, the world.
But, it has seemed to some philosophers, I only apprehend the world mediately, through my representations of it, either in ideas or in words
I do not have any access to the world as it is in itself.
(Those of you who have worked through the epistemology of the modern era will have a good understanding of the problem here.)
It seems as if I am unable to compare my words or my ideas to an independent world to decide whether there really is a correspondence between them.
The correspondence theory says that truth is a matching of words to the world, but I can only really know about one side of the equation.

In response to this problem with the correspondence theory of truth, some philosophers have adopted coherence theories.
According to a coherence theory, the truth of a sentence consists in its consistency with other beliefs we hold.
Different people apprehend the world in different ways, depending on their experiences, expectations, and background beliefs.
The coherentist despairs of any method of resolving these inconsistencies among people and their beliefs.

Imagine that I believe in a traditional, monotheistic God and that you do not.
1 will be true for me, since it coheres with my other beliefs.

1. God is omniscient.

In contrast, 1 will be false for you, since it conflicts with your other beliefs.
Since different people hold different beliefs, the coherence-truth of a sentence depends on the person who is considering the sentence.
Coherence theories thus collapse into relativism.

The correspondence and coherence theories of truth each provide a univocal analysis of 'truth'.
Insofar as they entail that there is a property called truth, they are both inflationary theories of truth.
Inflationary theories are distinguished from deflationary theories of truth.
Deflationary theories of truth were mainly developed in the last century, and are often called minimalist theories.
Deflationary theories have many proponents, all of whom have different ways of understanding and explaining deflationism.
Deflationists are united in the belief that there is no essence to truth, no single reduction of truth to a specific property, like correspondence or consistency.

Some deflationists claim that truth is just a device for simplifying long conjunctions.
If you said a lot of smart things at the party, I could list them all.
Or, I could just say:

2. Everything you said last night was true.

In 2, 'true' is eliminable by a long set of sentences listing all of what you said last night.
Abbreviations like 2 are, according to the deflationist, the central purpose of 'truth'.

Otherwise, 'truth' is merely a redundant term.
Indeed, deflationism is often called a redundancy theory of truth: to say that snow is white is true is just to say, redundantly, that snow is white.

Both inflationists and deflationists agree that a minimal condition for truth is what we call the T-schema, or Convention T, following Tarski.

3. p is true iff x

In 3, 'p' is the name of any sentence, and x are the truth conditions of that sentence.

We can use the T-schema to specify the truth conditions for any sentence.
Here are some instances of the T-schema:

4. 'The cat is on the mat' is true iff the cat is on the mat.
5. '2+2=4' is true iff 2+2=4
6. 'Barack Obama is president' is true iff the husband of Michelle Obama and father of Sasha Obama and Malia Obama is head of the executive branch of the United States of America.

Note that, as in 6, the truth conditions need not be expressed in the same terms as the sentence on the left.
We can even use a different language for the sentence and for its truth conditions.
Consider:

7. 'El gato está en el alfombrilla' is true iff the cat is on the mat.

Notice that you could understand the truth conditions of 7 without understanding the meaning of the Spanish sentence on the left side.

Inflationists and deflationists disagree about whether the T-schema is all there is to know about truth.
The inflationist believes that there are explanations of the concept of truth inherent in the truth conditions on the right side of the T-schema.
For the correspondence theorist, 'the cat is on the mat' is true because there is a cat, which corresponds to 'the cat', and there is a mat, which corresponds to 'the mat', and there is a relation, being on, which the cat and the mat satisfy, or in which they stand.
All other instances of the T-schema will have similar explanations in terms of the correspondence of words to worlds.

The deflationist, in contrast, believes that the T-schema is all there is to know about truth, and that there is no single kind of explanation of why all sentences are true.
'Truth' varies in application.
Again, according to the deflationist, we do not even need 'true' in our language.
It's just a handy tool.

I've discussed the deflationist and the coherence theorist mainly just to put Tarski's theory of truth, which Benacerraf describes as the only viable candidate, in perspective.

IV. Tarski and the Liar

In the early twentieth century, truth had gotten a terrible reputation, in large part due to the liar paradox, L, which leads to contradiction.

L: L is false

As we have discussed, contradictions are unacceptable in traditional, or classical, formal systems because a contradiction entails anything.
This property of classical systems is called explosion.

Explosion		
	1. P • ~P	
	2. P	1, Simp
	3. P ∨ Q	2, Add
	4. ~P	1, Com, Simp
	5. Q	3, 4, DS
QED		

To see how the liar leads to a contradiction, consider applying the T-schema (3) to L.
On the left side, we'll put 'L' for 'p'.
On the right side, we'll put 'L is false' for x.

TL L is true iff L is false.

We can translate this sentence into predicate logic by taking a constant, say 'p', to stand for the sentence L, and introducing a truth predicate, 'T'.
We also have to assume a bivalent logic, taking 'P is true' to be the negation of 'P is false'.
(There is a strengthened liar which persists even in three-valued logics).

1. Tp ≡ ~Tp	From the T-schema and the definition of P
2. (Tp ⊃ ~Tp) • (~Tp ⊃ Tp)	1, Equiv
3. ~Tp ⊃ Tp	2, Com, Simp
4. Tp ∨ Tp	3, Impl, DN
5. Tp	4, Taut
6. Tp ⊃ ~Tp	2, Simp
7. ~Tp ∨ ~Tp	6, Impl
8. ~Tp	7, Taut
9. Tp • ~Tp	5, 8, Conj

Tilt!

Our natural language contains the word 'true', as a predicate.
If we include a truth predicate in our formal language, we can construct the liar sentence.
If we can construct the liar sentence, we can formulate an explicit contradiction.
Contradictions explode.
Everything is derivable.
But, we know that not every sentence is true.
So, if we include a truth predicate in our formal language, our formal language will not be sound.

The excitement surrounding the new logic of the early twentieth century included hopes that all human knowledge could be represented by formal languages.
Since contradictions lead to explosion, and formal languages in which the paradoxes are representable lead to contradictions, it became seen as essential to avoid formalizing the notion of truth as in TL.
Since formal languages were seen as the locus of all of our knowledge, especially before Gödel's theorems, it seemed that truth was just not a legitimate term, not something that we could know.
The bad reputation of truth explains, at least in part, the interest of many philosophers in the relativism of coherence truth.
All recent work on truth, whether deflationary or inflationary, owes its origins to Tarski, who, in the 1930s, showed how to rehabilitate the concept of truth within formalized languages, how to avoid explosion without giving up on a formalized notion of truth.

The liar sentence L, like Russell's paradox, is self-referential.
Russell developed the theory of types in such a way as to prevent impredicative definitions, definitions which refer to themselves.
He relied on a vicious circle principle to eliminate such definitions.

> VC "Whatever involves *all* of a collection must not be one of that collection"; or, conversely: "If, provided a certain collection had a total, it would have members only definable in terms of that total, then the said collection has no total" (Whitehead and Russell, *Principia Mathematica*, Chapter II, p 37).

Tarski's theory of truth similarly proscribes self-reference.
Tarski rehabilitates the old Aristotelian view of truth by segregating object language from metalanguage.
The object language is the language in which we are working.
The meta-language is a language in which we can talk about the object language.
He banishes semantic terms from the object language, thus never allowing 'true' to apply to sentences which contain semantic terms, like 'false'.
But he allows semantic terms in the meta-language, as long as they apply only to terms of the object language.
We can construct theories of truth for the object language in the meta-language.

To be slightly more specific, imagine an object language that does not contain the word 'true'.
We can talk about the object language in a meta-language, as long as we have names for all of the sentences of the object language.
Furthermore, we can add a predicate to the meta-language, 'True', which applies to some sentences of the object language.
We can then partition the sentences of the object language into two classes: the true and the false.
Instances of the T-schema are sentences of the meta-language which we can use to characterize truth for the object language.

Deleting just the liar from the object language might appear arbitrary and ad hoc.
Tarski claims that the paradoxes show that all uses of the term 'true', and related semantic terms, are illegitimate within any object language.
We can not, on pain of contradiction, construct a truth predicate for a language within that language.
As long as we ascend to a meta-language, we can construct a truth predicate for any object language.
To determine which sentences of an object language are true and which are false, we have to examine the truth conditions as given on the right-hand side of instances of the T-schema.

While the sentences themselves are elements of the object language, the truth conditions are, technically, written in the meta-language.

The key to Tarski's solution to the liar paradox is that sentences like L are ill-formed, since they include 'false' in the object language.
When I want to use a sentence like 2, he claims, I implicitly ascend to a meta-language to do so.
In a meta-language, I can also construct sentences like the important 8.

> 8. All consequences of true sentences are true.

Sentences like 8 are fundamental to metalogic and model theory, fields that Tarski, following Hilbert, more or less created.
To formulate a theory of correspondence truth, Tarski's T-schema is not, by itself, sufficient.
As Hartry Field shows in a classic paper "Tarski's Theory of Truth," in order to use the T-schema as a definition of truth, we need to supplement it with some kind of account of why we choose certain sentences to be true and not others.
We can understand the truth conditions in 7 without understanding the Spanish sentence on the left.
To capture truth, it is not enough just to list the true and false sentences of a language.
We want to analyze the component parts of the Spanish expressions, and how they interact to form true or false sentences.
The T-schema, by itself, does not provide that kind of explanation.
Again, for 'the cat on the mat' to be true, 'the cat' must refer to a specific cat, 'the mat' must refer to a specific mat, and 'is on' must refer to the relation of being on, and the cat must be on the mat.
In order to know that 'the cat on the mat' is true, we have to know all of that, and not merely just the appropriate instance of the T-schema.

Tarski's construction thus reduces 'truth' to other semantic notions, like reference, that we need in order to interpret our sentences.
The semantic conception of truth depends on our metaphysics, on determining what there is and how our terms refer to it.

It is this use of Tarski's theory in a broader semantic account that Benacerraf believes the only viable theory of truth.

> I take it that we have only one such account: Tarski's, and that its essential feature is to define truth in terms of reference (or satisfaction) on the basis of a particular kind of syntactico-semantic analysis of the language, and thus that any putative analysis of mathematical truth must be an analysis of a concept which si a truth concept at least in Tarski's sense (Benacerraf 667).

V. Benacerraf's Problem

Most succinctly, Benacerraf's problem is that it seems impossible to match our epistemic capabilities with standard semantics, where semantics is taken to include theories of truth and reference.
The platonist mathematics that underlines standard interpretations of mathematical language seems incompatible with our epistemic capacities.

Benacerraf formulates his problem using two conditions on our theoretical commitments, on our overall view of the world.
The first condition is that we adopt a standard, uniform theory of truth: Tarski's.
Consider Benacerraf's sentences 1 and 2.

> B1. There are at least three large cities older than New York.
> B2. There are at least three perfect numbers greater than 17.

In first-order language, they have the following structures.

> B1'. $(\exists x)(\exists y)(\exists z)(Lx \bullet Ly \bullet Lz \bullet Oxn \bullet Oyn \bullet Ozn)$
> B2'. $(\exists x)(\exists y)(\exists z)(Px \bullet Py \bullet Pz \bullet Gxn \bullet Gyn \bullet Gzn)$

More abstractly, if less formally, they each have the form of Benacerraf's 3.

> B3. There are at least three FGs that bear R to a.

On a standard semantics, the truth of an existential sentence depends on whether there are objects in the domain of quantification that can substitute for the variables in the sentence so that the properties ascribed to those objects hold.
As Quine says, and we shall see, "To be is to be the value of a variable."
Standard semantic theory requires objects to satisfy the sentences.
Moreover, a standard semantic theory says that two sentences with the same grammatical structure are to be analyzed in the same way.
Their content might differ, but the structure should not depend on the content.

> One consequence of...the standard view is that logical relations are subject to uniform treatment: they are invariant with subject matter. Indeed, they help define the concept of "subject matter." The same rules of inference may be used and their use accounted for by the same theory which provides us with our ordinary account of inference, thus avoiding a double standard (Benacerraf 670).

Thus, just as we need cities to satisfy B1, we require mathematical objects to satisfy B2.
The standard view is platonist, since sentences like 'there is a number between 4 and 6' have the same grammar as sentences like 'there is a chair between the desk and the door'.
Both sentences say that there is an object.
The best way to understand the mathematical objects is as the platonist does.

Benacerraf's second condition is that we should have epistemic access to the objects to which our theory refers.
We have to have some account of how we can know about the things we think exist, which seems to be absent in the case of platonistic entities.
Benacerraf claims that the best theory of knowledge is CTK.
We need some sort of account of our connection to the objects of our theories, and CTK explains knowledge of causal laws.
CTK says, as we saw, that for me to know a claim, there must be a causal connection between me and the things to which the claim refers.
Further, Benacerraf holds, as a corollary, that the theory of reference is also causal, that my connection to

any object to which a term I know refers must be based in some causal link between me and the object.

The problem for standard semantics, then, is that we lack causal connection to the mathematical objects required to interpret sentences like B2.
Standard semantics posits objects outside the causal realm.

> If, for example, numbers are the kinds of entities they are normally taken to be, then the connection between the truth conditions for the statements of number theory and any relevant events connected with the people who are supposed to have mathematical knowledge cannot be made out. It will be impossible to account for how anyone knows any properly number-theoretical propositions (Benacerraf 673).

VI. Field's Reformulation

While Tarski's theory may not be all there is to say about truth, it remains the central, most viable theory of truth.
In contrast, CTK was a passing fad; as Field says, no one believes CTK anymore.
I mentioned the problems for CTK arising from our visit to fake barn country.

Field reformulates Benacerraf's challenge without CTK.

> The way to understand Benacerraf's challenge, I think, is not as a challenge to our ability to *justify* our mathematical beliefs, but as a challenge to our ability to *explain the reliability* of these beliefs... Benacerraf's challenge...is to provide an account of the mechanisms that explain how our beliefs about these remote entities can so well reflect the facts about them. The idea is that *if it appears in principle impossible to explain this*, then that tends to *undermine* the belief in mathematical entities, *despite* whatever reason we might have for believing in them (Field 25-6).

Whatever epistemic theory we have, whether it is CTK or another, we have to explain why our beliefs are reliable.
We have to explain why it is the case that if mathematicians believe p, then p.
Thus Field invokes a more-recent, and currently acceptable epistemic theory: reliabilism.
Reliabilism says that to be known, a belief must originate in reliable cognitive processes or faculties.

Reliabilism is a fairly abstract constraint on epistemology.
It doesn't claim that any particular processes (like sense perception) are reliable.
That's a job for psychology, I suppose.
It merely claims that the processes that underlie my knowledge must be reliable, as opposed to the processes in Gettier-style cases, or in fake-barn-country cases.
Any process can be acceptable, including mathematical intuition, as long as it can be shown to be reliable.

While Benacerraf's challenge is widely taken to be the locus of contemporary debates about mathematical epistemology, almost everyone these days takes it in Field's broader form.
The challenge is to find a way to bring together our theories of truth and reference, which seem to demand mathematical objects, with our best epistemic theories, which seem to debar knowledge of mathematical objects.

VII. The Gödel Solution

Benacerraf discusses and dismisses two attempts to solve the problem: from Gödel and from a group he calls combinatorialists, and which includes Hilbert, Brouwer, and Wittgenstein and the conventionalists.
Gödel, as we have seen, posited a faculty of mathematical intuition that could allow us to see mathematical truths.
Gödel's approach is inconsistent with the causal theory of knowledge.
But, as we have seen, Benacerraf never needed CTK anyway.
Benacerraf finds Gödel's view troubling for reasons independent of CTK.

> What troubles me is that without an account of *how* the axioms "force themselves upon us as being true," the analogy with sense perception and physical science is without much content. For what is missing is *precisely* what my second principle demands: an account of the link between our cognitive faculties and the objects known (Benacerraf 674).

Field dismisses Gödel's view, arguing that we lack reasons to believe that mathematical intuition is reliable.

> Someone *could* try to explain the reliability of these initially plausible mathematical judgments by saying that we have a special faculty of mathematical intuition that allows us direct access to the mathematical realm. I take it though that this is a desperate move... (Field 28)

Field asks us to imagine someone who claims to know precisely what is happening in a remote Nepalese village.
We wouldn't believe that person unless we were given some mechanism to explain that knowledge (say a webcam).
The Gödel-platonist owes us an explanation of the reliability of mathematical beliefs, not merely the assertion that they are reliable.

VIII. The Combinatorial Solution

Benacerraf's problem arises from the incompatibility of standard semantics with our best epistemology.
Gödel's view violates Benacerraf's epistemic constraint.
Another option is to give up standard semantics, as the views that Benacerraf calls combinatorial views do.
Benacerraf uses 'combinatorial' to refer to Hilbert's program, formalism, intuitionism, and conventionalism, because they construe mathematical truth as depending on the manipulation (or combination) of objects other than traditional mathematical objects.
Hilbert takes ideal mathematical terms to refer to inscriptions.
The strict formalist takes all mathematical terms to refer to their inscriptions.
The intuitionist invokes mental constructs as the referents of mathematical terms, as does Hilbert, for his real terms.
The conventionalist makes numerical terms refer to nothing at all.[1]

---

[1] Brown (156) reads Wittgenstein as a platonist, though not a realist (i.e. as an object platonist but not a sentence platonist). So, Brown's Wittgenstein admits of the standard semantics.

These views are united by their invocation of the manipulation of objects accessible to humans in their accounts of mathematics.
None of them accept a standard mathematical semantics.
They do not take '5' as referring to a five.

Against the combinatorialists, Benacerraf invokes and extends Quine's argument against conventionalism in logic.
Quine had argued, against Carnap's view, that we need logic in order to adopt a framework, so that the basic principles of a framework can not be merely conventional.
Benacerraf argues that a theory of truth demands a theory of reference.
We have not only to partition the set of statements of a theory into two classes.
We also have to know what those two classes are, and why we put some terms in one class and some in the other.
As long as we are using standard semantics, we need a standard theory of reference for mathematical terms.
If we appeal to the provability of mathematical theorems, and to the manipulation of non-mathematical objects, we still don't know why the axioms are true.
If we want a semantics which tells us why proofs yield truths, we must have a standard interpretation of the terms used.

> Although it may be a truth condition of certain number-theoretic propositions that they be derivable from certain axioms according to certain rules, *that* this is a truth condition must also follow from the account of *truth* if the condition referred to is to help connect truth and knowledge, if it is by their proofs that we know mathematical truths (Benacerraf 673).

The combinatorialists provide access to the objects of mathematics only by changing the subject.
They save their epistemology by abandoning their theory of truth.

> Motivated by epistemological considerations, they come up with truth conditions whose satisfaction or nonsatisfaction mere mortals can ascertain; but the price they pay is their inability to connect these so-called "truth conditions" with the truth of the propositions for which they are conditions (Benacerraf 678)

IX. Ways Out

Field argues that Benacerraf's argument, relying on an inference to the best explanation, is not a knock-down argument against platonism.
A knock-down argument would provide some explanation of why it is impossible to explain the reliability of beliefs about platonistic entities.
Non-existence proofs are difficult.

Field thus takes, as many do, the Benacerraf argument to be a challenge to the platonist.
But, it is really a problem for both platonists and anti-platonists.
The platonist has an access problem.
The anti-platonist has a semantics/truth problem.

One way to resolve Benacerraf's dilemma is to develop a Gödel-style epistemology.

That's the route taken by Jerrold Katz, and one that I favor.
Another way is to pursue the conventionalist account for mathematics, though perhaps not for logic.
Hartry Field defends fictionalism about mathematics, denying that mathematical sentences are true, but holding that they have standard truth conditions.

Lastly, and most importantly for us, the platonist can try to dissolve the access argument by appealing to a completely new way to determine ontology.
This is Quine's route, and the topic for the rest of the semester, in one form or other.
Quine's argument for the existence of mathematical objects is called the indispensability argument.
The indispensabilist argues that we have reliable beliefs about the physical world, and that reliability transfers to the mathematical objects.
Thus, the indispensabilist can try to answer Field's version of the Benacerraf dilemma by appealing to the reliability of our beliefs about empirical science.